

Statistical properties of planar Voronoi tessellations

H.J. Hilhorst

Laboratoire de Physique Théorique, Bâtiment 210, Univ Paris-Sud and CNRS, 91405 Orsay, France

Received: date / Revised version: date

Abstract. I present a concise review of advances realized over the past three years on planar Poisson-Voronoi tessellations. These encompass new analytic results, a new Monte Carlo method, and application to experimental data.

PACS. PACS 02.50.-r Probability theory, stochastic processes, and statistics – PACS 45.70.Qj Pattern formation – PACS 87.18.-h Multicellular phenomena

1 Introduction

In this talk I will concisely review a coherent collection of new results on planar Voronoi cells obtained over the last three years [1,2,3,4]. In Fig. 1 you see a *Voronoi tessellation*: the set of dots is given and I will refer to them as ‘centers’ or ‘point particles.’ Perpendicular bisectors have been constructed to the line segments connecting nearby particles. The bisectors meet in trivalent vertices (unless there is accidental degeneracy) and partition the plane into convex polygons called *Voronoi cells*.

In physics Voronoi cells have, roughly, two broad classes of applications. (i) They may directly model cellular struc-

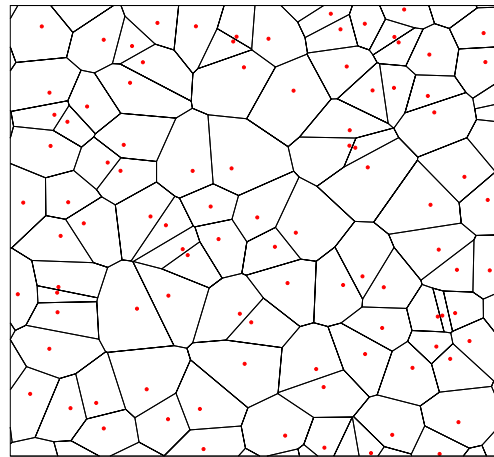


Fig. 1. Voronoi tessellation of a set of point particles.

tures, whether occurring naturally or synthesized; examples are biological tissue or soap froths. (ii) They may serve as a tool of analysis. For example, in a dense con-

figuration of real physical particles they define lattice defects; and if you decide to study your favorite theory on a lattice of randomly located spatial sites, the Voronoi construction is the natural way to define nearest-neighbor relations between these sites. Longer lists of applications may be found in the literature [5,6,2].

When the point particles are distributed independently and uniformly – which I will henceforth stipulate –, the resulting tessellation is called a *Poisson-Voronoi* tessellation. Real systems will of course most often exhibit deviations from this simple mathematical model: soap froths evolve with time and their cells are not constructed around a center; real particles have a minimum distance of approach; *etc.* However, the Poisson-Voronoi tessellation, being the simplest one to study, is to cellular systems what the Ising model is to magnetism.

Over the last 50 years, starting with the work of Meijering [7], many analytical properties of the Voronoi cell have been determined. These include the statistics of the perimeter segments, of the angles at the vertices, the cell area, and so on (an overview is given in Ref.[5]). However, the Voronoi cell's most prominent property, *viz.* its probability p_n of being n -sided, is extremely difficult to access analytically. The basic expression for p_n is a $2n$ -dimensional integral on the positions \mathbf{R}_m ($m = 1, \dots, n$) of n point particles neighboring a central one placed in the origin. A coupling between the \mathbf{R}_m arises from the condition that the resulting cell be n -sided. Hence the

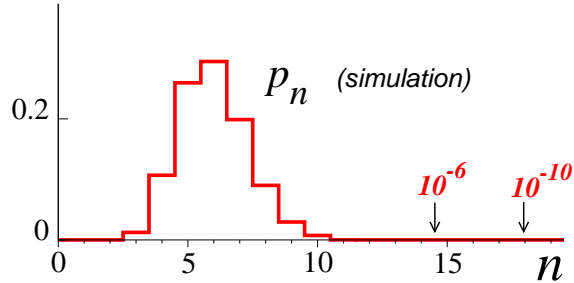


Fig. 2. Histogram of the sidedness probability p_n .

expression for p_n is on the same footing as a partition function of n interacting particles in two dimensions. No exact evaluation is known, but Monte Carlo simulations have led to the histogram of Fig. 2. It shows that p_n peaks at $n = 6$ and decreases very rapidly for larger n . It is as small as 10^{-6} for $n \approx 15$ and 10^{-10} for $n \approx 18$. Drouffe and Itzykson (DI) [8,9] devised a Monte Carlo algorithm that yielded p_n with a single-digit accuracy up to $n \approx 25$. The asymptotic decay of p_n for large n has been a subject of speculations. Many authors fitted p_n to a decaying exponential or stretched exponential, often multiplied by a power. DI proposed $p_n \sim n^{-\alpha n}$ with $\alpha \approx 2$.

2 Mathematics

A new line of research began with the mathematical challenge of *obtaining the analytic asymptotic expression of p_n for $n \rightarrow \infty$* . The calculation, although involving only classical tools of analysis, is of considerable complexity [1, 2]. I present briefly its main ideas.

A key role is played by the angles ξ_m and η_ℓ defined in Fig. 3. They define the cell up to a radial scale factor.

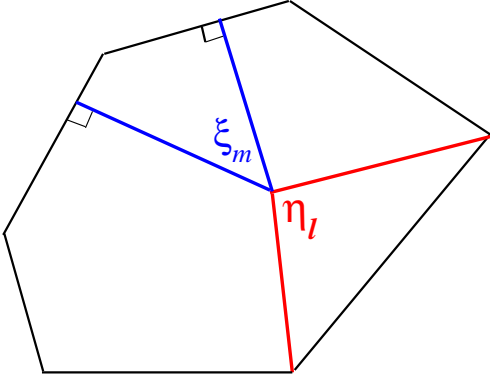


Fig. 3. The angles ξ_m and η_ℓ (with $m, \ell = 1, \dots, n$) that define the n -sided Voronoi cell up to a radial scale factor.

The obvious sum rules $\sum_{m=1}^n \xi_m = 2\pi$ and $\sum_{\ell=1}^n \eta_\ell = 2\pi$ represent infinitely weak constraints in the limit $n \rightarrow \infty$.

It appears that in that limit, to zeroth order in n^{-1} , the angles ξ_m and η_ℓ become a set of $2n$ independent variables, the distributions $P_1^{(0)}(n\xi_m/2\pi)$ and $P_2^{(0)}(n\eta_\ell/2\pi)$ being given by

$$P_1^{(0)}(x) = 4x e^{-2x}, \quad P_2^{(0)}(y) = e^{-y}, \quad x, y > 0. \quad (1)$$

This independence is nontrivial: the ξ_m and η_ℓ overlap, so how can they be uncorrelated? This property is best initially introduced as a hypothesis, to be confirmed self-consistently later by the solution procedure.

The independence of the angles can be shown to imply that as $n \rightarrow \infty$ the shape of the n -sided cell approaches a circle with probability 1. From an analysis of the radial integral in the expression for p_n one concludes that (for point particle density λ) this circle is of the radius $R_c(n) = \sqrt{n/4\pi\lambda}$.

After writing p_n in terms of the variables of integration ξ_m and η_ℓ , one can express it in the factorized form

$$p_n = C_n p_n^{(0)}, \quad p_n^{(0)} = \frac{2(8\pi^2)^{n-1}}{(2n)!}, \quad (2)$$

where $p_n^{(0)}$ results from integration with respect to the $P_1^{(0)}$ and $P_2^{(0)}$. Hence the problem of finding p_n has been replaced with that of finding C_n . This problem involves the $\mathcal{O}(n^{-1/2})$ and $\mathcal{O}(n^{-1})$ corrections to the infinite- n behavior. It appears that for C_n one can set up a perturbation expansion which shows that

$$C_n = C[1 + \mathcal{O}(n^{-1})], \quad n \rightarrow \infty, \quad (3)$$

where C can be expressed elegantly as

$$C = \prod_{q=1}^{\infty} \left(1 - \frac{1}{q^2} + \frac{4}{q^4}\right)^{-1} = 0.344\,347\dots \quad (4)$$

This infinite product is in fact the ‘elastic’ configurational partition function of the cell perimeter. The factor of index q in (4) stems from deviations from circularity having a wavelength $2\pi R_c/q$ along the perimeter.

There is an interesting corollary. It says that there is a continuum limit in which the cell perimeter, expressed as a function $R(\phi)$ of the polar angle ϕ , satisfies

$$\frac{d^2 R(\phi)}{d\phi^2} = L(\phi), \quad (5)$$

where $L(\phi)$ is Gaussian noise which is white to order zero but has colored order n^{-1} corrections. The solution of (5), under appropriate conditions pertaining to the average of $R(\phi)$ and to its periodicity [2], produces the full functional probability distribution $\mathcal{P}[R(\phi)]$ of an arbitrary perimeter

$R(\phi)$. Hence the asymptotic determination of p_n leads to a complete understanding of the behavior of the large n -sided cell.

Finally, the mathematical methods employed here are applicable to other problems that have arisen in mathematics and in physics. I mention two of them.

The Crofton problem. Let a plane be traversed by intersecting straight lines, distributed randomly and uniformly. This is another way of partitioning it into convex cells. The ‘Crofton problem’ is the question of determining the probability p_n^{Cr} for the cell containing the origin to be n -sided. This was done in Ref. [10], again in the limit of large n .

The problem of extremal points. Let n points be distributed randomly and uniformly in the unit disk. Then can one determine the probability p_n^{ext} that the convex envelope of this set is an n -sided polygon? Work on this problem is in progress.

3 Monte Carlo method

Attempts to estimate p_n by Monte Carlo simulation have been numerous (see Ref. [5]). Straightforward methods starting from a random set of centers suffer from the fact that many-sided cells are very rare. More sophisticated methods like those of Ref. [8] aim at directly generating n -sided cells for an n fixed in advance.

The derivation of the asymptotic result (3)-(4) for C_n has as its starting point a *non*asymptotic expression of

the form

$$C_n = \langle W_n \Theta_n \rangle, \quad (6)$$

valid for all $n = 3, 4, \dots$. Here $\langle \dots \rangle$ is an average with respect to the $P_1^{(0)}$ and $P_2^{(0)}$. Furthermore W_n is a weight that can be expressed [4], through a series of equations, in terms of the sets of angles $\{\xi_m\}$ and $\{\eta_\ell\}$, and Θ_n is a projector: $\Theta_n = 1$ if a certain condition on the angles is satisfied, and $\Theta_n = 0$ if not. This condition requires the n points \mathbf{R}_m all to be located such that they contribute a nonzero segment to the perimeter. All previous ‘fixed n ’ Monte Carlo methods described in the literature run into such an acceptance criterion. In all proposed cases, the projector rejects an ever larger fraction of configurations as n gets larger: this is the phenomenon of *attrition*, well-known (and deplored) in many Monte Carlo studies.

The particular split (2) that I made into a zeroth order problem and a remainder now appears to be the right one [4]. If the ξ_m and η_ℓ are chosen from the zeroth order distributions (1), then *the problem of attrition in the large- n limit is eliminated*: as n grows, the fraction of accepted configurations tends to unity! In practice, less than 1% of the configurations is rejected for $n \geq 20$, and less than 0.01% for $n \geq 40$.

Two things then become possible.

(i) To accurately determine the sidedness probabilities p_n for arbitrarily large n . Tables with values of p_n to at least four decimal places are given in Ref. [4] in the range $3 \leq n \leq 1600$. For high n these probabilities be-

come unphysically small: one has $p_{100} = 5.269 \times 10^{-188}$ and $p_{1000} = 6.36 \times 10^{-3841}$ (*sic!*), but I will show below which rewards can be gained from studying them.

(ii) To Monte Carlo generate typical n -sided cells for arbitrary *a priori* given n , together with their ‘natural’ environment of other cells. This is done as follows. Angles ξ_m and η_ℓ are randomly picked from the zeroth order distribution (but subject to the sum rules). If they pass the acceptance criterion, the cell perimeter is constructed. By radial scaling the cell radius is given its most probable value R_c . The positions of the central point particle and its first neighbors are now fixed. Next, all other point particles, in order that they do not interfere with the central Voronoi cell, are excluded from the ‘fundamental domain’ associated with that cell, *i.e.* from the union of the n disks centered at the cell vertices and passing through the origin; however, those other point particles (which are second and higher order neighbors to the central one), occupy the remaining region of the plane uniformly and randomly. Hence, after Monte Carlo generation of their positions, the construction of the full Voronoi tessellation may be completed by any of the existing algorithms.

Fig. 4 shows a 96-sided Voronoi cell constructed in this way. Snapshots of cells with $n = 24, 48, 96$, and 1536 may be found in Ref. [4].

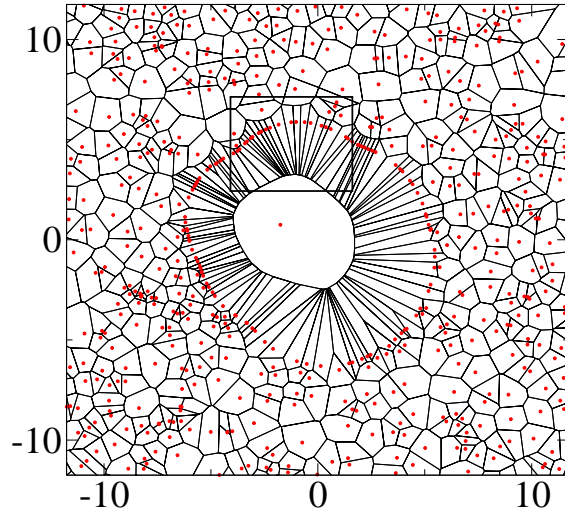


Fig. 4. Snapshot of a 96-sided Voronoi cell.

4 Aboav’s law

Aboav’s law, formulated in 1970 [11], holds that the neighbor of an n -sided cell has itself an average number m_n of sides given by

$$m_n = a + bn^{-1}, \quad (7)$$

where a and b are positive constants. Eq. (7) used to be, and is still often, formulated as $nm_n = an + b$, and is therefore called the ‘linear law’. Since m_n decreases with n , it says that “*many-sided cells tend to have few-sided neighbors, and conversely.*” Two-cell correlations have been measured in a large variety of experimental systems and found in agreement with Aboav’s law for parameters in the range $4.5 \lesssim a \lesssim 5.3$ (in this context 5 has been called a ‘magic number’) and $5.7 \lesssim b \lesssim 8.5$.

The first one to present heuristic arguments for the validity of (7) was Weaire [12], whence the alternative name ‘Aboav-Weaire law.’ Other ‘proofs’ of this law (see

Ref. [3] for some references) appeal to mean field approximations or proceed by maximizing a hypothesized entropy functional. While some workers consider Aboav's law as no more than a linear approximation to some unknown curve, others have attributed to it a more fundamental status, as is clear from numerous statements in the literature: *"In this paper we present an alternative derivation of the Aboav-Weaire law"* [13], or *"In all known natural and man-made structures, it is found empirically that nm_n is linearly related to n "* [14], and many others.

Now it has been known for over twenty years that Aboav's law does *not* hold exactly for the Poisson-Voronoi tessellation. This is borne out by Fig. 5: numerical simulation shows unambiguously that the m_n curve has a very small but distinct downward curvature. It is so small that, if also present in nature, most experimental data would not detect it.

In face of this, the defenders of Aboav's law as a basic truth hold that the Poisson-Voronoi tessellation, because of being constructed around random centers, does not correspond to any true physical structure. In real cellular structures, so they argue, any initial randomness has always undergone a relaxation process (for example to relieve internal stress); and this then would have induced the validity of Aboav's law.

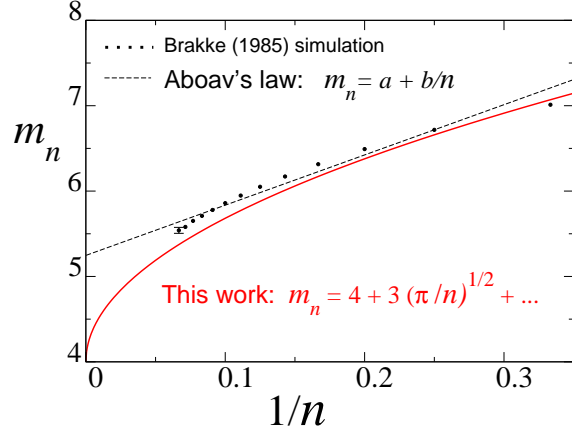


Fig. 5. Monte Carlo data, Aboav's law, and this work's asymptotic expression (8) for m_n .

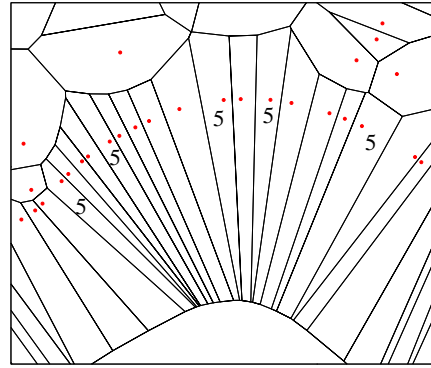


Fig. 6. Zoom onto the box in Fig. 4 showing some first-neighbor cells to the central cell.

4.1 Aboav's law and the Poisson-Voronoi tessellation

Before returning to real physical systems, let us first see what the present theory, further developed in Ref. [3], has to say about the validity of Aboav's law for the Poisson-Voronoi tessellation. The picture of Fig. 6 is useful for the argument. It shows a phenomenon that becomes ever more pronounced as n gets larger (whence the need to consider very large n ; see [4] for a picture with $n = 1536$): the first neighbors of the central cell become more and more elon-

gated while their width goes down to order $1/\sqrt{n}$. Since the second neighbors of the central cell have dimensions of order unity (for unit particle density), the geometry dictates that most first neighbors must become four-sided. The fraction of them that is five-sided is only of order $1/\sqrt{n}$; this is because five-sidedness occurs only when a first-order neighbor is adjacent to two second-order neighbors. In Fig. 6 all five-sided cells have been marked.

It follows that m_n is equal to 4 plus $\mathcal{O}(n^{-1/2})$ corrections. The latter require a nontrivial calculation [3] and one finds

$$m_n = 4 + 3\sqrt{\pi/n} + \dots, \quad n \rightarrow \infty. \quad (8)$$

In Fig. 5 the Monte Carlo data for m_n (due to Brakke [15]) are compared to the first two terms of the asymptotic expansion (8), as well as to the best linear two-parameter fit of type (7). Clearly the asymptotic theory that I developed explains for the first time why m_n must be curved; moreover, it predicts correctly the order of magnitude of this curvature [3].

4.2 Aboav's law and experimental systems

Under suitable conditions, polystyrene latex spheres of diameter $\approx 1\mu m$, when trapped at a water/air interface, will on a time scale of hours undergo a process of diffusion limited colloidal aggregation (DLCA). The experiment was performed by Earnshaw and coworkers [16, 17, 18] in the 1990's and was simulated very recently by

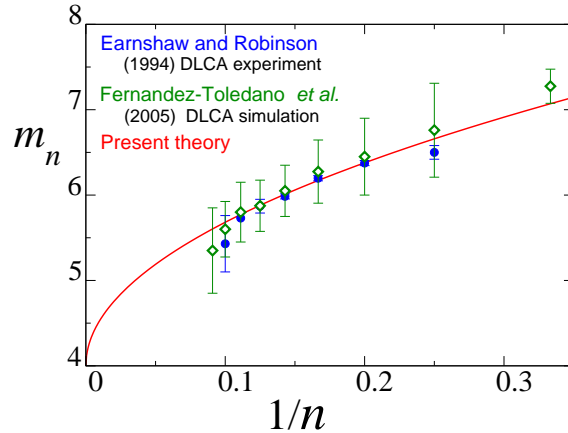


Fig. 7. Experimental and simulational m_n data compared to the asymptotic Poisson-Voronoi curve (8).

Fernández-Toledano *et al.* [19]. Snapshots of the system were taken at regular time intervals and the center of gravity of each cluster of particles was determined. Then the Voronoi cells belonging to this set of centers were constructed and m_n was obtained. Fig. 7 shows the experimental DLCA data (dots) of Ref. [16] taken after 60 seconds, as well as the corresponding simulation data (diamonds) of Ref. [19]. Also shown is the same theoretical curve (solid line) as in Fig. 5. In view of the experimental and numerical error bars, this figure does not *prove* that the DLCA system is of the Poisson-Voronoi type. I believe however that, as an explanation for these data, the present zero-parameter curvature-predicting theory is preferable to the linear two-parameter fit (7).

A final point is worth discussion. In the experiment, the clusters whose centers of gravity serve as the ‘point particles’ of the Voronoi construction, have some finite diameter d . Is this a problem? The experiment is done,

typically, at an area coverage of about 10% [16], and hence at a cluster number density λ such that $\pi d^2 \lambda / 4 \approx 0.1$. In the Poisson-Voronoi model, for a point particle density λ , the typical distance between two first neighbors of an n -sided cell is $\ell_n \equiv 2\pi(2R_c)/n = \sqrt{4\pi/n\lambda}$. The results of the present asymptotic theory can be expected to be applicable only for $n \lesssim n^*$, where n^* is a sidedness such that $\ell_{n^*} \approx d$. The point particle/cluster density λ drops out of this relation and we find $n^* \sim 100$. Taking into account the possibly fractal structure of the clusters may lower this estimate of n^* , but the n values of Fig. 7 are less than n^* by what seems a safe margin. Crossover between the ideal Poisson-Voronoi behavior and a finite-diameter theory was briefly discussed in Ref. [3].

5 Conclusion

What we learn from this is that in other experimental situations, too, as the range and the precision of the data increase, a curved m_n law may well be expected and must indeed be looked for.

And, on a more general level, that provided you ask the right questions, *the patterns induced by random points in a plane are challenging!*

References

1. H.J. Hilhorst, *J. Stat. Mech.* (2005) L02003.
2. H.J. Hilhorst, *J. Stat. Mech.* (2005) P09005.
3. H.J. Hilhorst, *J. Phys. A* **39** (2006) 7227.
4. H.J. Hilhorst, *J. Phys. A* **40** (2007) 2615.
5. A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, second edition (John Wiley & Sons Ltd., Cichester, 2000).
6. N. Rivier in: *Disorder and Granular Media*, eds. D. Bideau and A. Hansen, Elsevier (1993).
7. J.L. Meijering, *Philips Research Reports* **8** (1953) 270.
8. J.M. Drouffe and C. Itzykson, *Nuclear Physics* **B235** [FS11] (1984) 45.
9. C. Itzykson and J.M. Drouffe, *Statistical field theory* (Cambridge University Press, Cambridge, 1989), Vol. 2, ch. 11.
10. P. Calka and H.J. Hilhorst, *in prepration*.
11. D.A. Aboav, *Metallography* **3** (1970) 383.
12. D. Weaire, *Metallography* **7** (1974) 157.
13. S.F. Edwards and K.D. Pithia, *Physica A* **205** (1994) 577.
14. B. Dubertret, N. Rivier, and M.A. Peshkin, *J. Phys. A* **31** (1998) 879.
15. K.A. Brakke (1986) *unpublished*. Available on <http://www.susqu.edu.brakke/aux/downloads/200.pdf>.
16. J.C. Earnshaw and D.J. Robinson, *Phys. Rev. Lett.* **72** (1994) 3682.
17. J.C. Earnshaw and D.J. Robinson, *Physica A* **214** (1995) 23.
18. J.C. Earnshaw, M.B.J. Harrison, and D.J. Robinson, *Phys. Rev. E* **53** (1996) 6155.
19. J.C. Fernández-Toledano, A. Moncha-Jordá, F. Martínez-López, A.E. González, and R. Hidalgo-Álvarez, *Phys. Rev. E* **71** (2005) 041401.